(In)sécurité de la Voix sur IP (VoIP)

Nicolas FISCHBACH

COLT Telecom/Sécurité.Org
nico@{colt.net,securite.org}

1 Introduction

Jusqu'à récemment, la voix sur IP était plutôt la technologie d'une minorité de "férus du net" car elle permet de téléphoner à moindre coût via l'Internet, la voix restant une forme de communication bien plus conviviale que le courrier électronique ou les formes de messageries instantanées. Avec l'évolution de l'Internet, le déploiement de réseaux privés virtuels, l'introduction de la qualité de service dans les réseaux, l'arrivée des PBX IP, la disponibilité de postes téléphoniques intégrant des fonctionnalités de plus en plus avancées ainsi que l'opportunité de réduction des coûts, l'intérêt que suscite la voix sur IP pour l'entreprise et pour les utilisateurs est grandissant jusqu'à devenir un projet stratégique. Tout d'abord les différents protocoles ainsi que leurs évolutions récentes seront présentés, puis les éléments qui composent une solution VoIP ainsi que les différentes formes d'attaques et les moyens de s'en prémunir ou de réduire les risques, et enfin pour terminer, une comparaison de la sécurité de la voix sur IP par rapport au réseau téléphonique classique et au réseau GSM.

2 Les différents protocoles

La voix sur IP est une description relativement générique et ne définit pas une liste exclusive de protocoles, ni d'équipements. Avant d'étudier les aspects sécurité nous allons présenter les différents protocoles, dont certains sont en train de devenir obsolètes, et d'autres émergents.

2.1 Le protocole H.323

H.323 est le premier protocole développé pour permettre des communications multi-médias. SIP son " concurrent ". H.323 est relativement complexe et SIP tente de simplifier les échanges en utilisant une sémantique proche de HTTP.

H.235 définit des mécanismes de sécurité.

2.2 SIP

SIP (*Session Initiation Protocol*) est le standard IETF pour la signalisation (établissement, terminaison, redirection, relayage, etc) de communications multimédias interactives. Ce protocole est de nos jours celui qui est déployé couramment. Le format est proche d'une adresse de messagerie : *sip :nico@securite.org*

SIMPLE (*SIP for Instant Messaging and Presence Leveraging Extensions*) est une extension de SIP dont le but est de supporter les messageries instantanées.

Le projet PROTOS s'est intéressé à SIP, et comme pour SNMP, a trouvé un nombre important d'implémentations qui ne passaient pas le " batch " de tests sans se planter ou redémarrer. Cela concerne aussi bien les téléphones, que les relais SIP et les équipements de sécurité.

2.3 RTP

RTP (*Real Time Transport Protocol*) encode, transporte et décode la voix.

2.4 MEGACO/H.248

MEGACO (*MEdia GAteway COntrol Protocol*) est un protocole commun à l'IETF et l'ITU-T qui joue le rôle d'interface entre des passerelles (par exemple entre la passerelle qui gère la signalisation et celle qui gère le flux voix). C'est une évolution de MGCP (*Media Gateway Control Protocol*). A la différence de H.323 et de SIP qui reposent sur une architecture distribuée, MEGAGO et MGCP sont orientés client/serveur (ie. centralisé).

3 Les protocoles secondaires

3.1 DNS

Le service DNS est utilisé pour fournir des services d'annuaire et de localisation.

3.2 TFTP et HTTP

Ces deux protocoles sont utilisés par les téléphones et différents autres éléments pour télécharger leur configuration.

3.3 ENUM

ENUM permet de lier des adresses SIP via DNS aux numéros de téléphone au format E.164

4 Architecture d'une solution VoIP

Après cette liste de protocoles et d'acronymes, qui bien que longue est loin d'être exhaustive, nous présentons les différents éléments qui composent l'architecture d'une solution VoIP.

4.1 Le réseau

Le réseau interconnecte les différents équipements qui participent à une solution de voix sur IP. A la différence du réseau téléphonique, la signalisation et la voix sont transportés sur le même réseau partagé. Comme pour IPsec, des solutions alternatives ont dû être développées pour gérer les contraintes introduites par la traduction d'adresses (NAT).

4.2 Le téléphone

Le téléphone peut se présenter sous deux formes : soit un téléphone " classique ", soit un téléphone " logiciel " qui s'éxécute sur l'ordinateur de l'utilisateur. En terminologie SIP c'est un UA (*User Agent*).

4.3 Les systèmes

La liste des protocoles usités dans les solutions de téléphonie et de voix sur IP est conséquente. Il en va de même de celle des systèmes.

4.4 Serveurs SIP : proxy, redirect et registrar

Le " *SIP proxy* " joue le rôle de relais et fait suivre une requête SIP au prochain serveur. Le " *SIP redirect server* " renvoit une réponse au client contenant le prochain serveur à contacter. Le " *SIP registrar* " enregistre le lien entre l'adresse IP et l'adresse SIP.

4.5 Le " Call Manager " / IP PBX

Le CM fournit des fonctions de base de gestion d'appel, des utilisateurs et des configurations, et également des fonctionnalités avancées comme la conférence, les boîtes vocales, etc. Il peut être vu comme un IP PBX. A ne pas confondre avec des PBX traditionnels (pas de VoIP) que l'on peut administrer à distance via une connexion TCP/IP (qui remplace la connexion locale ou de télé- maintenance via un modem).

4.6 Les passerelles

Une passerelle s'occupe des échanges entre le réseau voix sur IP et le réseau téléphonique classique. Cela comprend la conversion de la voix et de la signalisation.

5 Les équipements de sécurité

5.1 Pare-feu

Bon nombre de pare-feux se limitent à gérer l'ouverture de ports en fonction des communications et n'inspectent pas les flux (au niveau protocolaire). De plus cet élément additionnel risque d'introduire un délai ainsi qu'une gigue, c'est pourquoi ils sont absents dans bien des déploiements.

5.2 IDS

Il n'est pas très courant de trouver des outils de détection d'intrusion pour des solutions de voix sur IP. La quantité de faux positifs dus à l'observation du flux RTP pourrait être plus que conséquente. Bien qu'elle permette de détecter des dénis de service par exemple, la détection d'intrusion se ramène souvent à de la détection de fraude.

6 Les attaques et les parades

6.1 Les types d'attaques

Déni de service Les attaques par déni de service se retrouvent sous plusieurs formes. Les plus classiques sont celles qui visent à utiliser toute la bande passante disponible ou abuser de problèmes intrinsèques à TCP/IP.

Dans le cadre d'une solution VoIP bien des éléments peuvent être attaqués : le téléphone, le réseau, le système d'exploitation, l'application, etc. Autant un déni de service sur l'Internet peut être filtré avec des mécanismes et des techniques plus ou moins avancées, autant celui à l'encontre d'une communication sera difficile à traiter et aura un impact direct sur les possibilités de communications.

Par exemple un nombre trop important de messages SIP INVITE ou de simples messages ICMP peuvent créeer une situation de déni de service.

Interception L'interception d'une communication peut être l'oeuvre d'un " criminel informatique " ou des autorités. En effet, l'interception légale de communications via un réseau de données (LI – *Lafwul Intercept on packet switched networks*, standard ETSI pour l'Europe et CALEA - *Communications Assistance for Law Enforcement Act* par exemple) est un sujet de discussion dans bon nombre de pays. Reste à voir comment sera traitée la voix sur IP par rapport aux communications téléphoniques classiques où l'opérateur est quasiment obligé d'être impliqué.

Les techniques bien connues d'écoute de réseau (" sniffing " ou de l'homme du milieu (" man in the middle " s'appliquent à l'interception. En revanche, contrairement à une attaque MITM contre un protocole comme telnet ou HTTP où il est possible de modifier le contenu à la volée, cela s'avère plus contraignant et plus facilement détectable avec de la voix encodée.

Le protocole RTP transporte la voix encodée, sans aucun chiffrement, ce qui rend l'interception relativement triviale.

" Call-ID " Le service de présentation du numéro de l'appelant est devenu, pour bon nombre d'abonnés le facteur principal de prise ou de rejet d'appel, tout particulièrement depuis l'essor des téléphones portables.

A la différence du téléphone classique où le numéro est lié à la ligne physique et qu'un voisin indélicat ne peut prendre ou écouter vos appels qu'à condition de mettre des pinces crocodiles sur votre ligne, ou écouter la bonne fréquence si

vous utilisez un téléphone sans-fil, cela est différent dans une solution VoIP. En effet, la mobilité qu'apporte la téléphonie sur IP introduit également un problème d'authentification de l'utilisateur : celui-ci doit se souvenir de son nom/numéro d'utilisateur et de son mot de passe, que quelqu'un pourrait lui " voler ", et donc s'enregistrer à sa place ou de manière concurrente.

Il est également relativement simple de manipuler l'identifiant de l'appelant. L'impact n'est pas très important, sauf dans le cas ou le CLID est utilisé pour authentifier l'appelant et autoriser l'accès à une ressource (boîte vocale, appels internationaux ou numéros spéciaux, etc...).

Non-répudiation et fraude La fraude la plus connue est l'accès gratuit ou à un côut réduit à des services à valeur ajoutée ou des appels internationaux.

La compromission de serveurs Les serveurs jouent un rôle important dans une solution de voix sur IP, et même s'il n'est pas forcément possible d'intercepter un appel si un serveur est compromis, il est souvent possible de récupérer des CDRs (*Call Detail Records*) qui contiennent toutes les traces des appels effectués. En revanche la compromission d'une passerelle entre le réseau VoIP et le réseau téléphonique classique permet d'écouter de manière transparente les appels, même s'ils sont chiffrés du côté VoIP (SRTP).

6.2 L'architecture sécurisée

Les systèmes et le réseau En fonction des options de déploiement choisies, il est probable que la majorité des systèmes doivent être accessibles depuis " partout ". Il convient donc de sécuriser ces éléments comme tout serveur, et dans la mesure du possible de mettre en place une solution avec des pare-feux.

La séparation entre le trafic voix et la signalisation peut se faire au niveau du réseau à l'aide de VLANs.

Déni de service Les différents éléments qui composent une architecture de type SIP, comme le relais, sont livrés par certains vendeurs avec des mécanismes de détection de déni de service (comme l'envoi massif de messages INVITE).

Le mécanisme le plus important est la qualité de service (QoS) Sans déploiement de bout-en-bout, et tout particulièrement sur les réseaux locaux où ce n'est pas très commun, il est possible pour n'importe qui de générer un déni de service. La bande passante disponible est également un facteur, mais la QoS et la gestion des files d'attentes est le point clé.

La voix sur IP n'implique pas que le trafic utilise l'Internet comme média, un déni de service sur un réseau privé est plus difficile à mettre en oeuvre.

Des solutions de filtrage de contenu ou d'analyse anti-virus commencent à être disponibles et utilisent SIP pour permettre l'inspection, un peu à l'image de protocoles comme OPSEC.

Interception La seule solution pour limiter l'interception est l'usage de mécanismes cryptographiques pour les flux de signalisation et de données (voix encodée). Ce chiffrement devrait être de bout en bout et il est important d'en évaluer les impacts. Le plus important étant celui sur la qualité de la communication à cause du délai supplémentaire introduit. Comme dans toute solution cryptographique, l'authentification forte des parties impliquées dans les échanges est un élément, sinon les attaques classiques de MITM peuvent s'appliquer.

Comme dans toute solution où les données sont chiffrées, elles doivent être déchiffrées. Cela nous ramène à la problématique du poste client. Autant il était impossible de mettre un cheval de Troie ou une porte dérobée sur le téléphones de nos grands-parents, autant il est aujourd'hui possible de " patcher " quasiment n'importe quel téléphone... sans parler des téléphones logiciels (" soft phones ").

La version " S " de SIP, SIPS, reposant sur TLS (*Transport Layer Security*) sur TCP permet de chiffrer les échanges SIP qui contiennent, par exemple, le nom d'utilisateur, le mot de passe ainsi que le numéro appelé.

Une nouvelle version du protocole RTP, S-RTP intègre des mécanismes de chiffrement.

Non-répudiation et fraude La fraude peut être limitée par la configuration correcte de la passerelle VoIP vers RTC (Réseau Téléphonique Commuté) ainsi qu'une gestion des droits/classes de service par utilisateur. Les passerelles doivent être configurées pour éviter qu'un utilisateur se connecte directement sans passer par le relais SIP.

7 Les réseaux de téléphonie " classiques "

A titre de comparaison, nous allons discuter, sans rentrer dans les détails, de la sécurité des réseaux téléphoniques plus classiques.

7.1 PSTN/POTS

Le réseau téléphonique filiaire (PSTN/POTS – *Public Switched Telephone Network/Plain Old Telephone System*) transporte voix et données. A la différence d'un réseau VoIP où les équipements qui " gèrent " les communications sont souvent dans le même réseau et accessibles, cela n'est pas le cas dans le RTC (SS7 et le " switch " par exemple).

7.2 GSM

Le réseau cellulaire (GSM – *Global System for Mobile Communications*) à la différence d'un réseau sans fil de type 802.11 ne permet pas d'écouter facilement une communication. En revanche les communications ne sont chiffrées qu'entre le téléphone de l'utilisateur et la station à laquelle il est rattaché. Ce n'est pas un chiffrement de bout-en-bout entre l'appelant et l'appelé, mais des solutions

existe qui emploient par exemple le mode données pour transporter de la voix chiffrée.

Un changement majeur est apparu ces dernières années avec le déploiement des réseaux de troisième génération (GPRS et UTMS) : les téléphones ne sont plus des grille-pains mais sont livrés avec un système d'exploitation, Java, une pile TCP/IP, etc. Et de plus sont connectés en permanence au réseau et disposent d'une adresse IP.

8 Conclusion

Du fait que tous les équipements " parlent " IP et que bien souvent l'Internet soit le média de transport, et également par manque d'éducation des utilisateurs il est clair le nombre d'attaques à l'encontre de solutions voix sur IP est plus large que celle ä l'encontre du RTC ou GSM. Peut-être également parce que bon nombre d'attaques sur l'Internet sont relativement anonymes et qu'il n'est pas facile de les tracer.

A propos

Nicolas FISCHBACH est Senior Manager chez COLT Telecom et dirige l'équipe sécurité au sein du département européen d'ingénierie IP. Il est également co-fondateur de Sécurité.Org, un site web francophone dédié à la sécurité informatique, d'eXperts, un groupe informel de spécialistes sécurité ainsi que du chapitre français du Honeynet Project. Nicolas participe à de nombreuses conférences (BlackHat Briefings, CanSecWest, Defcon, JSSI, Eurosec, NANOG, Cisco Systems, RIPE, SwiNOG, Libre Software Meeting, etc), publie des articles (MISC) et donne des cours dans différentes écoles et universités (HEC, Université de Genève, ITIN, etc). Pour plus d'informations : `http://www.securite.org/nico/`